

(19)日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開2000-39970

(P2000-39970A)

(43)公開日 平成12年2月8日(2000.2.8)

(51)Int.Cl. ⁷	識別記号	F I	テ-マ-ト*(参考)
G 0 6 F 3/06	3 0 5	G 0 6 F 3/06	3 0 5 C 5 B 0 0 1
	5 4 0		5 4 0 5 B 0 1 8
11/10	3 3 0	11/10	3 3 0 L 5 B 0 6 5
12/16	3 2 0	12/16	3 2 0 L
G 1 1 B 20/18	5 2 0	G 1 1 B 20/18	5 2 0 A

審査請求 有 請求項の数 4 O L (全 5 頁) 最終頁に続く

(21)出願番号 特願平10-209658

(22)出願日 平成10年7月24日(1998.7.24)

(71)出願人 000192545

神戸日本電気ソフトウェア株式会社

兵庫県神戸市西区高塚台5丁目3番1号

(72)発明者 森本 節夫

兵庫県神戸市西区高塚台5丁目3番1号

神戸日本電気ソフトウェア株式会社内

(74)代理人 100065385

弁理士 山下 稔平

Fターム(参考) 5B001 AA02 AB02 AD04

5B018 GA02 HA16 HA35 MA14

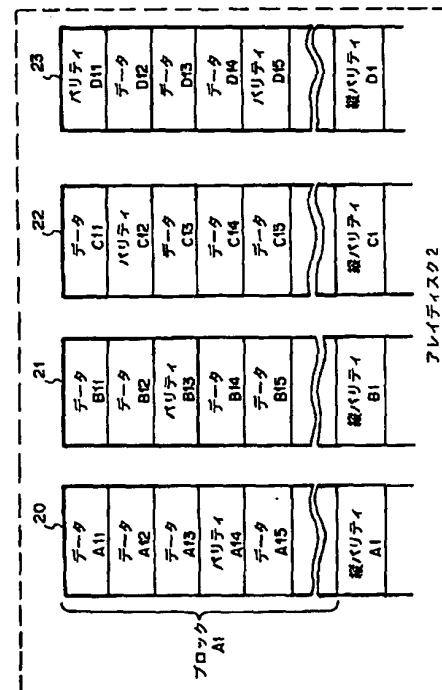
5B065 BA01 CA30 EA02 EA13 EA24

(54)【発明の名称】 ディスクアレイシステムの二重障害防止制御方式

(57)【要約】

【課題】 ディスク1台が故障し代替ディスクにデータを再構築する場合に残りのディスクでメディアエラーが発生してもデータが正常に補間できるシステムを提供する。

【解決手段】 一部のディスクに他ディスクに書き込まれるデータを基に生成される第1パリティが書き込まれ、同一ディスク内のデータを基に生成される第2パリティが該同一のディスクに書き込まれるアレイディスクと、第1パリティを生成する第1パリティ生成手段と、第2パリティを生成する第2パリティ生成手段と、アレイディスクの一部のディスクに障害が発生したときに、他のディスクのデータ乃至第1パリティを用いて障害が発生したディスクのデータを訂正する第1訂正手段と、アレイディスクのあるディスクのデータ乃至第2パリティにエラーが発生したときにエラー以外の部分からエラーを訂正する第2訂正手段とを備える。



【特許請求の範囲】

【請求項1】 一部のディスクに他のディスクに書き込まれるデータを基に生成される第1の-parityが書き込まれ、同一のディスク内のデータを基に生成される第2の-parityが該同一のディスクに書き込まれるアレイドiskと、

前記第1の-parityを生成する第1の-parity生成手段と、

前記第2の-parityを生成する第2の-parity生成手段と、

前記アレイドiskの一部のディスクに障害が発生したときに、他のディスクのデータ乃至前記第1の-parityを用いて障害が発生したディスクのデータを訂正する第1の訂正手段と、

前記アレイドiskのあるディスクのデータ乃至前記第2の-parityにエラーが発生したときにエラー以外の部分からエラーを訂正する第2の訂正手段と、

を備えることを特徴とするディスクアレイスistemsの二重障害防止制御方式。

【請求項2】 請求項1に記載のディスクアレイスistemsの二重障害防止制御方式において、前記アレイドiskは訂正されたデータ乃至第1及び第2の-parityを書き込むための予備のディスクを備えることを特徴とするディスクアレイスistemsの二重障害防止制御方式。

【請求項3】 請求項1又は2に記載のディスクアレイスistemsの二重障害防止制御方式において、前記第1の-parityが書き込まれるディスクが前記第1の-parityを含む符号語毎に異なることを特徴とするディスクアレイスistemsの二重障害防止制御方式。

【請求項4】 請求項1乃至3のいずれか1項に記載のディスクアレイスistemsの二重障害防止制御方式において、前記第2の-parityを記憶する記憶手段を備え、前記第2の-parity生成手段は、現在のデータと新たにアレイドiskに書き込まれるデータとの差分と前記記憶手段内の前記第2の-parityを基に前記記憶手段内の第2の-parityを更新することを特徴とするディスクアレイスistemsの二重障害防止制御方式。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明はディスクアレイスistemsの二重障害防止制御方式に関する。

【0002】

【従来の技術】 従来、ディスクアレイスistemsにはデータベースなどの貴重なデータが保存されておりディスクの故障に伴うデータの消失は許されない状況で使用されている。

【0003】 なお、関連する従来技術として、特開昭59-110098号公報に記載の「データ記憶装置の誤り訂正装置」、特開平6-230903号公報に記載の「ディスクアレイス装置の障害回復方法、およびディスク

アレイス装置」、特開平6-289998号公報に記載の「ディスクアレイス装置」、特開平7-28608号公報に記載の「安全性の高いディスクアレイスの改良」及び特開平9-167120号公報に記載の「記憶装置の誤り訂正装置」がある。

【0004】

【発明が解決しようとする課題】 ディスクアレイスistemsも正常な状態で運用されている場合は、ディスクの故障及びメディアエラーに対しては安全なシステムであるが、ディスクが1台故障した場合は逆に非常に障害が発生しやすいシステムになりうる。

【0005】 本発明は、ディスク1台が故障し代替ディスクにデータを再構築する場合において残りのディスクでメディアエラーが発生しても、データが正常に補間できるシステムを提供することを目的とする。

【0006】

【課題を解決するための手段】 本発明によるディスクアレイスistemsの二重障害防止制御方式は、一部のディスクに他のディスクに書き込まれるデータを基に生成される第1の-parityが書き込まれ、同一のディスク内のデータを基に生成される第2の-parityが該同一のディスクに書き込まれるアレイドiskと、前記第1の-parityを生成する第1の-parity生成手段と、前記第2の-parityを生成する第2の-parity生成手段と、前記アレイドiskの一部のディスクに障害が発生したときに、他のディスクのデータ乃至前記第1の-parityを用いて障害が発生したディスクのデータを訂正する第1の訂正手段と、前記アレイドiskのあるディスクのデータ乃至前記第2の-parityにエラーが発生したときにエラー以外の部分からエラーを訂正する第2の訂正手段と、を備えることを特徴とする。

【0007】 また、本発明によるディスクアレイスistemsの二重障害防止制御方式は、上記のディスクアレイスistemsの二重障害防止制御方式において、前記アレイドiskは訂正されたデータ乃至第1及び第2の-parityを書き込むための予備のディスクを備えることを特徴とする。

【0008】 更に、本発明によるディスクアレイスistemsの二重障害防止制御方式は、上記のディスクアレイスistemsの二重障害防止制御方式において、前記第1の-parityが書き込まれるディスクが前記第1の-parityを含む符号語毎に異なることを特徴とする。

【0009】 更に、本発明によるディスクアレイスistemsの二重障害防止制御方式は、上記のディスクアレイスistemsの二重障害防止制御方式において、前記第2の-parityを記憶する記憶手段を備え、前記第2の-parity生成手段は、現在のデータと新たにアレイドiskに書き込まれるデータとの差分と前記記憶手段内の前記第2の-parityを基に前記記憶手段内の第2の-parityを更新することを特徴とする。

10

20

30

40

50

【0010】

【発明の実施の形態】本発明によるディスクアレイシステムの二重障害防止制御方式は、ディスクアレイシステムに冗長なデータを付加することによって二重障害を最小限にするものである。

【0011】図1において、アレイコントローラ1はアレイディスク2の初期化時、各ディスク20～23間のパリティに加え、図3の様に各ディスク20～23内にブロックを定義したときのブロック内の縦パリティを作成初期化する。この縦パリティはアレイコントローラ1及び各ディスク20～23内に保存する。CPU3からアレイコントローラ1へライトコマンドが発行されると、アレイコントローラ1はまずアレイディスク2から更新される前のデータを読み込む。次に各ディスク20～23間のパリティの計算と同時に、はじめに各ディスクから読みとったデータと書き込むデータの差分から縦パリティを作成し、その後データとパリティをアレイディスク2に書き込む。これによりディスクの故障によるアレイ障害時の復旧作業に中において、復旧に使用している他のディスクにメディアエラーが発生した時にこの縦パリティを使用しそのディスクのデータを復旧させることが可能であるため、アレイシステムの二重障害を最小限に抑えることができる。

【0012】図1を参照すると、アレイコントローラ1にはライト／リードの命令を行うCPU3、データを保持する数台のディスク群であるアレイディスク2が接続されている。

【0013】次に、アレイコントローラ1の具体的な構成及び、各ディスクの使用方法について説明する。

【0014】図2は、本発明の実施形態におけるアレイコントローラ1の構成例を示すブロック図である。図2において、アレイ制御回路10は中心的な回路であり各装置及び回路を制御する。アレイ制御回路10には、ディスクコントローラ11、縦パリティ保持メモリー12、不揮発性メモリー13が接続されている。縦パリティ保持メモリー12及び不揮発性メモリー13は、アレイディスク2へのアクセス量を低減させる目的に使用する。

【0015】図3はアレイディスク2の構成例である。アレイディスクの先頭に書かれているデータを例に説明すると、データA11、データB11、データC11が1組のデータであり、その組に対するパリティがパリティD11である。以下のデータもパリティのあるディスクを順序よく変えながら同様に構成してある。以上は、現在一般的に使用されているディスクアレイのデータの構成例であるが、さらに本実施形態においては各ディスク内だけでブロック（例：ブロックA1）を定義し、そのブロックに対するパリティ（例：縦パリティA1）もディスクに収納されている。

【0016】次に、図1～図3のシステムの動作につい

て、図を参照して説明する。

【0017】まず、本実施形態のディスクアレイシステムを使用するにあたってディスクの初期化を行う必要があるが、その際に、ディスク20～23間のパリティの生成、各ディスク20～23内のブロックの定義、及びブロック内の縦パリティの生成を行う。

【0018】ディスクアレイシステム運用にあたって、電源が投入されるとまずアレイコントローラ1は各ディスク20～23の縦パリティの全てを縦パリティ保持メモリー12から読み込む。

【0019】CPU3からアレイコントローラ1に対してデータライトの命令が発効されると、まずアレイ制御回路10はアレイディスク2からそのデータライトの対象となるエリアのデータを読み出す。またそのデータライトの対象となるエリアに対応するブロックの縦パリティも同時に縦パリティ保持メモリー12から読み出す。次に、更新される前のデータとライトするデータの差分と現在の縦パリティから新たな縦パリティを計算する。ディスク間のパリティも従来通りそのエリアのデータから作成し、その後データ及びディスク間のパリティをアレイディスク2に、縦パリティを縦パリティ保持メモリー12にそれぞれ書き込む。電源投入後ディスクアレイのデータが更新された場合は、不揮発性メモリー13にデータ更新のビットを立てる。これはディスクアレイ運用中に停電などが発生した場合、データと縦パリティが一致しないことを示すために用いる。CPU3からリード命令が発行された場合は、パリティ、縦パリティの参照及び変更はせず、アレイディスク2から読み込むだけである。

【0020】運用を終了する場合は、縦パリティをディスクに書き戻し不揮発性メモリー13のデータ更新ビットを下げて電源を切る。次回本システムを立ち上げた時に、不揮発性のメモリーのデータ更新のビットが立っていた場合は、運用前にディスクのデータを読み込み再度縦パリティを生成しその後運用する。不揮発性のメモリーのデータ更新のビットが立っていない場合は、システムを立ち上げた時に読み込む縦パリティをそのまま使用できる。

【0021】アレイディスク2のディスク20～23のうちのあるディスク2nが運用中故障した場合、残りの正常な全ディスクから代替ディスク24（不図示）に故障したディスク2nのデータを再構築するが、残りのディスクでメディアエラーが発生した場合、同一ディスク内のブロックのデータ及び縦パリティを使用してデータを復旧させ二重障害を防止する。

【0022】

【発明の効果】従来は、ディスクアレイシステムは構成するディスクの内1台が故障等により使用不可となると一般のディスク構成より非常に障害が発生する可能性が高い状態になる。任意のディスクが故障すると、故障し

10

20

30

40

50

たディスクのデータは残りのディスク間のパリティを使用し再生可能であるが、残りのディスクにおいてメディアエラーが発生した場合はデータを再生できなくなる。代替ディスクが用意されたシステムでは残りのディスクの全データから代替ディスク用のデータを再構成しなければならない。その残りのディスク内で1つのメディアエラーが発生しても代替ディスクへ書き込むデータの再生は不可能になる。このメディアエラーはディスクの物理的特性によりある確率で発生するものであるため回避する方法はない。

【0023】本発明により、縦パリティを付加することによりデータの再構築中に同一ディスク内で定義されたブロック内でメディアエラーが1つ発生しても、縦パリティからデータを補間できるので、二重障害につながる可能性は大幅に低減できる。

【0024】従って、本発明によれば、ディスクアレイシステムにおいて二重障害を最小限に抑えられるという効果が奏される。

【図面の簡単な説明】

【図1】本発明の実施形態によるディスクアレイシステムの二重障害防止制御方式の構成を示すブロック図である。

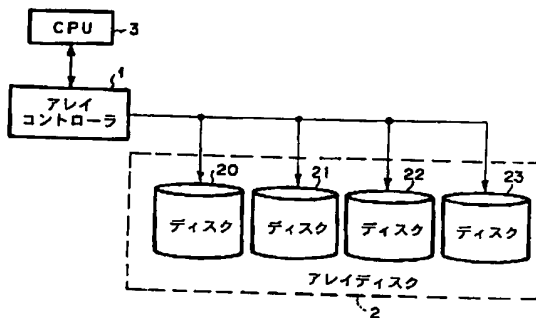
【図2】図1に示すアレイコントローラの構成を示すブロック図である。

【図3】図1に示すアレイディスク2の記録フォーマットを示す図である。

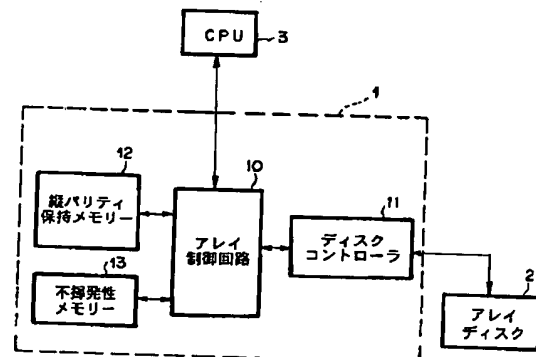
【符号の説明】

- 10 アレイコントローラ
- 2 アレイディスク
- 3 CPU
- 10 アレイ制御回路
- 11 ディスクコントローラ
- 12 縦パリティ保持メモリ
- 13 不揮発性メモリ
- 20～23 ディスク

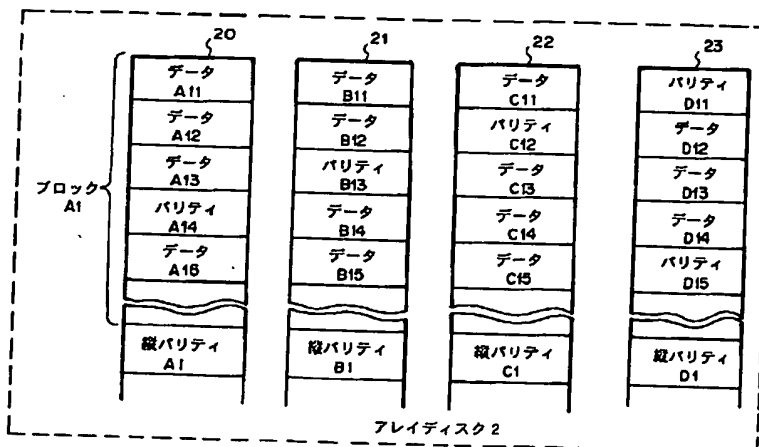
【図1】



【図2】



【図3】



(5)

特開2000-39970

フロントページの続き

(51)Int.Cl.⁷

G11B 20/18

識別記号

570

572

F I

G11B 20/18

テ-マコ-ト*(参考)

570Z

572F